Available online at www.sciencedirect.com
ScienceDirect
Fisheries Research 86 (2007) 188–200
FISHERIES RESEARCH
www.elsevier.com/locate/fishres

Monthly catch forecasting of anchovy *Engraulis ringens* in the north area of Chile: Non-linear univariate approach

Juan Carlos Gutiérrez-Estrada [a,*], Claudio Silva [b], Eleuterio Yáñez [b], Nibaldo Rodríguez [c], Inmaculada Pulido-Calvo [a]

[a] *Dpto. Ciencias Agroforestales, EPS, Campus La Rábida, Universidad de Huelva, 21819 Palos de la Frontera, Huelva, Spain*
[b] *Escuela de Ciencias del Mar, Facultad de Recursos Naturales, Pontificia Universidad Católica de Valparaíso, Casilla 1020, Valparaíso, Chile*
[c] *Escuela de Informática, Facultad de Ingeniería, Pontificia Universidad Católica de Valparaíso, Av. Brasil 2241, Valparaíso, Chile*

Received 5 February 2007; received in revised form 4 June 2007; accepted 7 June 2007

Abstract

In this study the performance of computational neural networks (CNNs) models to forecast 1-month ahead monthly anchovy catches in the north area of Chile considering only anchovy catches in previous months as inputs to the models was analysed. For that purpose several CNN approaches were implemented and compared: (a) typical autoregressive univariate CNN models; (b) a convolution process of the input variables to the CNN model; (c) recurrent neural networks (Elman model); (d) a hybrid methodology combining CNN and ARIMA models. The results obtained in two different external validation phases showed that CNN having inputs of anchovy catches of the 6 previous months hybridised with ARIMA(2,0,0) provided very accurate estimates of the monthly anchovy catches. For this model, the explained variance in the external validation fluctuated between 84% and 87%, the standard error of prediction (SEP, %) was lower than 31% and mean absolute error (MAE) was around 18,000 tonnes. Also, significant results were obtained with recurrent neural networks and seasonal hybrid CNN + ARIMA models. The strong correlation among estimated and observed anchovy catches in the external validation phases suggests that calibrated models captured the general trend of the historical data and therefore these models can be used to carry out an accuracy forecast in the context of a short-medium term time period.

Keywords: Computational neural network; Recurrent neural network; Elman model; ARIMA model; Hybrid model; Catch prediction

1. Introduction

The pelagic fish commonly named anchovy (*Engraulis ringens*, Jenyns 1842) is together with South American pilchard (*Sardinops sagax*, Jenyns 1842) one of the most important fishing resources in the north area of Chile. Both species have been caught in coastal waters since the beginning of the 1950s and actually both species support the 42% of total landing (about 2.1 million of tonnes per year). As a consequence of the anchovy population collapse, in 1983 this fishery in the north area of Chile is subjected to regulation during the most important reproductive periods. In this scenario, the forecasting of catches is a basic topic, because it plays a central role in management of stocks, preceding decision making (Makridakis et al., 1983). In fisheries management policy the main goal is to establish the applicable fishing effort in a concrete area during a known period keeping the stock replacements. To achieve this aim, it is necessary to predict uncontrollable events, such as possible abundance or biomass changes.

The application of statistical and mathematical tools to relevant data in order to obtain a quantitative understanding of the stocks status and quantitative predictions of stock reactions to alternative future regimes have been used for many years. In a general way, three of the major approaches presented in the literature to assess the biomass available involve models based on biological-fishery variables relationships, 'black box' approaches and methods that synthesise the two approaches mentioned previously.

A model based on biological-fishery variables relationships generally aims to formulate the biological-fisheries process in terms of each of its relevant component, such as catch-at-age

* Corresponding author. Tel.: +34 959217528; fax: +34 959217528.
E-mail addresses: juanc@uhu.es (J.C. Gutiérrez-Estrada), claudio.silva@ucv.cl (C. Silva), eyanez@ucv.cl (E. Yáñez), nibaldo.rodriguez@ucv.cl (N. Rodríguez), ipulido@uhu.es (I. Pulido-Calvo).

0165-7836/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2007.06.004

of several years, historical series of catch-effort data, length–weight relationship, parameters of the population growth model or time series of spawning stocks and recruitments. Thus, this type of model is often too complex and demanding in terms of data (Lassen and Medley, 2001). Also, the accuracy of the model predictions is generally affected by assumed components which are often highly dependent on the expertise and experience of the user concerning fishery behaviour (Hilborn and Walters, 1992; Shepherd, 1999).

In a 'black box' approach a model is applied to identify a direct mapping between inputs and outputs without detailed consideration about the internal structure of the biological and fishery processes. Although a fishery is a complex system and therefore it cannot be completely described in a simple form, there are many practical situations (such as catch real-time accurate predictions) in which a 'black box' approach may be preferred to not expend the time and effort required to develop, validate and implement a model based on biological-fishery variables.

It should be pointed out that data availability often determines the model choice. In fact, monthly anchovy catches can be easily obtained when compared with data of fish age, length–weight relationship, age-length key or parameters of the growth model, natural mortality or time series of spawning stocks and recruitment. Therefore, a 'black box' approach that operates only based on the first data set can be much more suitable for operational forecasting purposes than a model based on biological-fishery variables that also requires the latter set of measurements.

Artificial or computational neural networks (ANNs or CNNs) can be classified as 'black box' type models. A CNN is a non-linear mathematical structure capable of representing the complex non-linear processes that relate the inputs to the outputs of a system. CNNs models are being increasingly applied in many fields of science and engineering and usually provide highly satisfactory results. Some specific applications of CNN to management and planning of fisheries include the modeling of abundance, recruitment, stock biomass, distribution and catch of different fisheries (Komatsu et al., 1994; Chen and Ware, 1999; Huse and Gjosaeter, 1999; Fréon et al., 2003; Hardman-Mountford et al., 2003; Huse and Ottersen, 2003; Maravelias et al., 2003; Hyun et al., 2005).

In the time series forecasting issues past observations of one or more variables are collected and introduced as input data in a model that describes the underlying relationships among those variables and allows estimating future realizations of one of the same (Zhang, 2003). Recently, CNNs have been extensively applied to time series forecasting (Griñó, 1992; Prybutok et al., 2000; Belgrano et al., 2001; Zeng et al., 2001; Gutiérrez-Estrada et al., 2004; Pulido-Calvo and Portela, 2007; Velo-Suárez and Gutiérrez-Estrada, 2007), although few studies have been applied in fisheries sciences (Komatsu et al., 1994; Chen et al., 2000; Chen and Hare, 2006).

This paper evaluates the performance of feed forward CNN models trained with the Levenberg–Marquardt algorithm (Shepherd, 1997) for the purpose of anchovy catches time series forecasting (one-step monthly anchovy catch forecast model). In order to verify if more accurate CNN solutions could be achieved, three variations of the typical autoregressive univariate CNN models or additions to those models were tested, namely: (a) a convolution process of the input variables to the CNN model (De Vries and Principe, 1991); (b) recurrent neural networks (Elman model); (c) a hybrid methodology combining CNN and ARIMA models were applied to take advantage of the unique strength of both models in non-linear and linear modeling, respectively (Wedding and Cios, 1996; Hansen and Nelson, 1997; Zhang, 2003).

2. Methods

2.1. Data source

The anchovy catches have been obtained from three principal data bases: (a) the annual fishery statistics of the Fishery National Service of Chile (SERNAPESCA, 1978–2004); (b) the Agricultural and Livestock National Service of Chile (SAG, 1963–1977); (c) the data bases of the Sea Institute of Perú (www.imarpe.gob.pe). The final data set was composed by anchovy catches from January 1963 to December 2005.

2.2. Computational neural networks: general concepts

Computational neural networks (CNNs) are mathematical models inspired by the neural architecture of the human brain. CNNs can recognize patterns and learn from their interactions with the 'environment'. The most widely studied and used structures are multilayer feed forward networks (Rumelhart et al., 1986). A typical four-layer feed forward CNN has g, n, m and s nodes or neurons in the input, first hidden, second hidden and output layers, respectively [the notation of the neural network is (g,n,m,s)]. The parameters associated with each of the connections between nodes are called weights. All connections are 'feed forward'; that is, they allow information transfer only from an earlier layer to the next consecutive layers.

Each node j receives incoming signals from every node i in the previous layer. Associated with each incoming signal (x_i) there is a weight (W_{ji}). The effective incoming signal (I_j) to node j is the weighted sum of all the incoming signals, according to:

$$I_j = \sum_{i=1}^{g} x_i W_{ji} \tag{1}$$

The effective incoming signal, I_j, passes through an activation function (sometimes called a transfer function) to produce the outgoing signal (y_j) of the node j. In this study, the linear function (l) ($y_j = I_j$) was used in the output layer and the sigmoid non-linear function (s), in the hidden layers:

$$y_j = f(I_j) = \frac{1}{1 + \exp(-I_j)} \tag{2}$$

in which I_j can vary on the range $(-\infty, \infty)$, and y_j is bounded between zero and one. Because of the use of sigmoid functions in the CNN model, the values of the data variables must be normalized onto range [0, 1] before applying the CNN methodology.

To determine the set of weights a corrective-repetitive process called 'learning' or 'training' of the CNN is performed. This training helps to define the interconnections among neurons (weights), and it is accomplished by using both known inputs and outputs (training sets or patterns), and presenting these to the CNN in some ordered manner, adjusting the interconnection weights until the desired outputs are reached. The strength of these interconnections is adjusted using an error convergence technique so that a desired output will be produced for a given input. There are many training methods. In this work, a variation of back-propagation algorithm (Rumelhart et al., 1986), known as the Levenberg–Marquardt algorithm (Shepherd, 1997) was applied. This is a second-order non-linear optimization algorithm with a very fast convergence and it is recommended by several authors (Tan and van Cauwenberghe, 1999; Anctil and Rat, 2005).

Levenberg–Marquardt algorithm operates by assuming that the underlying function being modeled by the neural network is linear. Based on this assumption, the minimum of the objective function can be exactly determined in a single step. The computed minimum is tested, and if the error is lower than in the previous step, the algorithm moves the weights to the new point. This process is iteratively repeated on each generation. Since the linear assumption is ill-founded, it can easily lead Levenberg–Marquardt to test a point that is lower (perhaps substantial lower) than the current one. The most ingenious aspect of Levenberg–Marquardt is that the computation of the new point is actually a compromise between a step in the direction of steepest descent and the above-mentioned leap. Successful steps are accepted and lead to a strengthening of the linearity assumption (which is approximately true near a minimum). Unsuccessful steps are rejected and lead to a more cautious 'downhill' step. Thus, Levenberg–Marquardt continuously switches its approach and can make very fast progress.

Levenberg–Marquardt algorithm uses the following formula that is continuously updated:

$$\Delta W = -(Z^{\mathrm{T}}Z + \lambda I)^{-1} Z^{\mathrm{T}}\varepsilon \tag{3}$$

where ε is the vector of errors, Z the matrix of the partial derivatives of these errors with respect to the weights W, and I is the identity matrix. The first term of the second member of the Levenberg–Marquardt formula represents the linear assumption and the second, the gradient-descent step. The control parameter λ governs the relative influence of these two approaches. Each time Levenberg–Marquardt succeeds in lowering the error, it decreases the control parameter by a factor of 10, thus strengthening the linear assumption and attempting to jump directly to the minimum. Each time it fails to lower the error, it increases the control parameter by a factor of 10, giving more influence to the descent gradient step, and also making the step size smaller.

On the other hand, the Elman network is commonly a two-layer network with feedback from the first layer output to the first layer input. Theoretically, this recurrent connection allows the Elman network to both detect and generate time-varying patterns. This model has recurrent neurons in its hidden layer (recurrent layer), and normal neurons in its output layer (the same as a typical feed forward CNN). This combination is special because two-layer networks can approximate any function (with a finite number of discontinuities) with an arbitrary accuracy. The only requirement is that the hidden layer must have enough neurons. Therefore, more hidden neurons are needed as the function being fit increases in complexity (Elman, 1990).

Let epoch denotes the time period that encompasses all the performed iterations after all the patterns are displayed. In the study presented in this paper, the learning process was controlled by the method of internal validation. Over-training is probably the most common error in training neural networks. The best method of ensuring that over-training does not occur is to monitor periodically the sum square error for both the training data and the selected data for internal validation. It is normal for the sum square error for the training data to continue to decrease with training. However, this may be forcing the neural network to fit the noise in the training data. To avoid this problem, stop periodically the training, substitute the internal validation data for one epoch, and record the sum square error. When the sum square error of the internal validation data begins to increase, the training is stopped and the weights at the previous monitoring are selected for the external validation phase (Tsoukalas and Uhrig, 1997).

2.3. General procedure

The data set (anchovy catches from 1963 to 2005) was divided in two subsets: the first one was composed by anchovy catches from 1963 to 2004 and the second one was composed by anchovy catches of year 2005. In the first subset, the 60% of data (randomly selected) were used for the CNNs training or calibration, the 20% (randomly selected) were used for internal validation and the 20% remaining were used in the testing, generalisation or external validation. This external validation phase was denoted as external validation type I (EVI). The second subset (year 2005) was used to carry out a second external validation a data set belong to full annual cycle of anchovy catches. This second external validation phase was denoted as external validation type II (EVII).

The inputs number (number of lagged months considered) were determined by means the analysis of the partial autocorrelation function, autocorrelation function and spectral analysis of the data series. The number of nodes in the hidden layers was determined by trial and error. CNN architectures with 2 hidden layers and 5–20 hidden nodes (hidden layers topology used: 5*s*–5*s*; 10*s*–10*s*; 15*s*–15*s* and 20*s*–20*s*; where *s* is the sigmoid transfer function) were successively trained based on the calibration data set.

An inherent problem associated with CNNs is their tendency to get stuck in local minima. To alleviate this problem, the same CNN is trained more than once, starting with a random set of weights (Anctil and Rat, 2005). The best CNN is then selected as the one with the lowest error in the external validation phase. In this paper, a pool of 30 repetitions were carried out for each CNN architecture because this level of repetitions implies that the chosen model is among the best 14% of the distribution of all possible models at the 99% confidence level (Iyer and Rhinehart,

1999). This sensibility analysis allowed the determination of the best CNN architecture.

CNN models were implemented using STATISTICA 6.0. In the case of the recurrent neural network, hidden topologies of 5–55 neurons were proved. In the same way, 30 repetitions for each model were calibrated and validated. These models were calibrated using Neural Networks Toolbox of MATLAB 6.5.

2.4. Computational neural network with smoothing process of the input variables

The multilayer feed forward network is a static CNN architecture. To assess the temporal patterns, the CNN must have an appropriate memory to store past information. The simplest form of memory consists of a buffer that contains the δ most recent inputs, that is to say, that contains multiple copies of the significant input data at various time delays δ (the anchovy catches in previous months $t-1$, $t-2$, etc.). Another common methodology is to represent memory as a convolution of the input sequence of monthly catches x_i with a kernel or smooth function (De Vries and Principe, 1991). Additionally, an advantage that presents the application of a smooth function is the high-frequency noises removing (Fréon et al., 2003). In the carried out analysis, buffer/smooth function methodologies were implemented. In this work, the smoothing function used was:

$$Q'_t = \alpha Q_t + (1-\alpha)Q'_{t-1}, \quad \alpha = \frac{1}{3} \tag{4}$$

where Q'_t is the anchovy catches time series smoothed, Q_t the anchovy catch in the t instant, α a smoothed coefficient and Q'_{t-1} is the anchovy catch smoothed in the $t-1$ instant.

2.5. Computational neural network and ARIMA model: hybrid approach

ARIMA models and CNNs are often compared with no clear conclusions in terms of the relative forecasting performance superiority (Zhang, 2003). ARIMA models assume that a time series is a linear combination of its own past values and of current and past values of an error term (Box and Jenkins, 1976). In this paper, a hybrid approach to time series forecasting using both CNN and ARIMA models was analysed. The main idea of the combination of these two models was to use each model feature to capture different patterns in the data. The methodology implemented consists of two steps: (a) in the first one, a CNN is developed to model the anchovy catches and (b) in the second one, an ARIMA model is used to describe the residuals from the CNN model. The ARIMA model helps to interpret the unexplained variance by the CNN model.

The CNN and ARIMA combined model can be formulated as follows:

$$Q_t = f(Q_{t-1}, Q_{t-2}, \ldots, Q_{t-\delta}) + \phi^{-1}(B)\theta(B)\eta_t \tag{5}$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{6}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{7}$$

where f is a function determined by the neural network structure and connection weights; ϕ_j ($j=1,\ldots,p$) the weights of the ARIMA model associated with each previous observation ε_t (in the applications carried out, the residuals from the CNN model); θ_j ($j=1,\ldots,q$) the weights of the ARIMA model associated with each previous noise terms; B the backshift operator that assigns a value to a variable in the previous instant ($B\varepsilon_t = \varepsilon_{t-1}$ and $B^p\varepsilon_t = \varepsilon_{t-p}$); η_t is the noise term in instant t. Differencing transformation (d) is often applied to the data to remove the trend and to stabilize the mean before an ARIMA model can be fitted.

In the ARIMA(p,d,q) (P,D,Q)^S model, values of p and q varying from zero to six (with a unitary step), and values of d varying from zero to two (also with a unitary step) were tested. The values of p, d, and q that proved to be more appropriate according to the accuracy measures presented in Section 2.6 were then used. The parameters ϕ_j and θ_j were fixed by using function minimization procedures, so that the sum of squared residuals was minimized. The level of significance of these parameters was evaluated (acceptable if $p<0.05$). This approach was combined with the best CNN determined by the previous trainings. The S value is the seasonality and P, D, Q are seasonal terms.

Extraction of the periodical components of the time series can be easily done using the autocorrelation (ACF) and partial autocorrelation functions (PCAF) (Holton-Wilson and Keating, 1996) and Fourier transformation (FFT). The Fourier transformation for discrete time series is defined as (Park, 1998):

$$F(k) = \sum_{n=0}^{N-1} f(n)\, \mathrm{e}^{(-j2\pi kn/N)} \tag{8}$$

where $f(n)$ is the discrete time series, $F(k)$ the Fourier transform of $f(n)$ and N is the total number of months. From the real and imaginary parts of the Fourier transformation is possible to obtain the periodogram, which is defined as:

$$\text{periodogram} = (\text{real part } F(k))^2 (\text{imaginary part } F(k))^2 \frac{N}{2} \tag{9}$$

2.6. Measures of accuracy applied in the model validation phase

To assess the performance of the models during the validation phases several measures of accuracy were applied, as there is not a unique and more suitable performance evaluation test (Yapo et al., 1996; Legates and McCabe, 1999; Abrahart and See, 2000). The correlation between observed and predicted catches was expressed by means of the correlation coefficient R. The coefficient of determination (R^2) describes the proportion of the total variance in the observed data that can be explained by the model. Other measures of variances applied were the percent standard error of prediction (SEP) (Ventura et al., 1995), the coefficient of efficiency (E_2) (Nash and Sutcliffe, 1970; Kitanidis and Bras, 1980) and the average relative variance (ARV) (Griñó, 1992). These four estimators are unbiased estimators that are employed to see how far the model is able to explain the total variance of the data.

In addition, it is advisable to quantify the error in the same units of the variables. These measures, or absolute error mea-

sures, included the square root of the mean square error (RMSE) and the mean absolute error (MAE), given by:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(Q_t - \hat{Q}_t)^2}{N}} = \sqrt{\mathrm{MSE}},$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}|Q_t - \hat{Q}_t|}{N} \tag{10}$$

where Q_t is the observed anchovy catch at the time step t; $\hat{Q}_t$ the estimated anchovy catch at the same time step t; N is the total number of observations of the validation set.

The percent standard error of prediction, SEP, is defined by:

$$\mathrm{SEP} = \frac{100}{\bar{Q}}\mathrm{RMSE} \tag{11}$$

where $\bar{Q}$ is the average of the observed anchovy catches of the validation set. The principal advantage of SEP is its non-dimensionality that allows comparing in a same basis forecasts given by different models. The coefficient of efficiency E_2 and the average relative variance ARV are used to see how the model explains the total variance of the data and represent the 'proportion' of the variation of the observed data considered by the model. E_2 and ARV are given by:

$$E_2 = 1.0 - \frac{\sum_{i=1}^{N}|Q_t - \hat{Q}_t|^2}{\sum_{i=1}^{N}|Q_t - \bar{Q}|^2},$$

$$\mathrm{ARV} = \frac{\sum_{i=1}^{N}(Q_t - \hat{Q}_t)^2}{\sum_{i=1}^{N}(Q_t - \bar{Q})^2} = 1.0 - E_2 \tag{12}$$

The sensitivity to outliers due to the squaring of the difference terms is associated with E_2 or, equivalently, with ARV. A value of zero for E_2 indicates that the observed average $\bar{Q}$ is as good as predictor as the model, while negative values indicate that the observed average is a better predictor than the model (Legates and McCabe, 1999).

For a perfect match, the values of R^2 and of E_2 should be close to one and those of SEP and ARV close to zero.

Also the persistence index, PI, was used for the model performance evaluation (Kitanidis and Bras, 1980):

$$\mathrm{PI} = 1 - \frac{\sum_{i=1}^{N}(Q_t - \hat{Q}_t)^2}{\sum_{i=1}^{N}(Q_t - Q_{t-L})^2} \tag{13}$$

where Q_{t-L} is the observed catch at the time step $t-L$ and L is the lead-time. In the carried out applications L was set equal to one, since only 1-month ahead forecasts were performed. A PI value of one reflects a perfect adjustment between predicted and observed values, and a value of zero is equivalent to say that the model is not better than a naïve model, which always gives as prediction the previous observation. The PI is well designed for assessing forecasts, as the anchovy catches in previous months are the main CNN inputs. A negative PI value would mean that the model is degrading the original information, thus denoting a worse performance than the one of the naïve model (Anctil and Rat, 2005).

In this paper, the indexes Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were calculated to give the possibility to compare with values obtained by other authors. These indexes are given by (Qi and Zhang, 2001):

$$\mathrm{AIC} = \log(\mathrm{MSE}) + \frac{2m}{N}, \qquad \mathrm{BIC} = \log(\mathrm{MSE}) + \frac{m\log(N)}{N} \tag{14}$$

In the previous equations m is the number of parameters of the model. In these equations both first terms of the second members measure the goodness-of-fit of the model to the data while the second terms penalise the model parameters number. The number of model parameters has been considered as the number of neural network weights as in other works have been done (Qi and Zhang, 2001; Chen and Hare, 2006). But in any way, it would be necessary to indicate that the neural network weights have not the same features as the parameters of an equation obtained using standard statistical estimations techniques (like linear regressions).

For each measure of accuracy the benchmark of the worst permissible error was calculated. McLaughlin (1983) suggests that a naïve model determines the forecasting accuracy benchmark of any model. The basic naïve model, known as 'Naïve Forecast I' (NFI) is defined as the period of the next level will be the same as that of the preceding period. This way, if the forecasting model cannot do better than NFI, it should be rejected. In the case of AIC and BIC a value of $m = 1$ was considered.

3. Results

3.1. ACF, PCAF and FFT of anchovy catch time series

Fig. 1(a) shows the autocorrelation (ACF) and partial autocorrelation (PCAF) functions of the original anchovy catches time

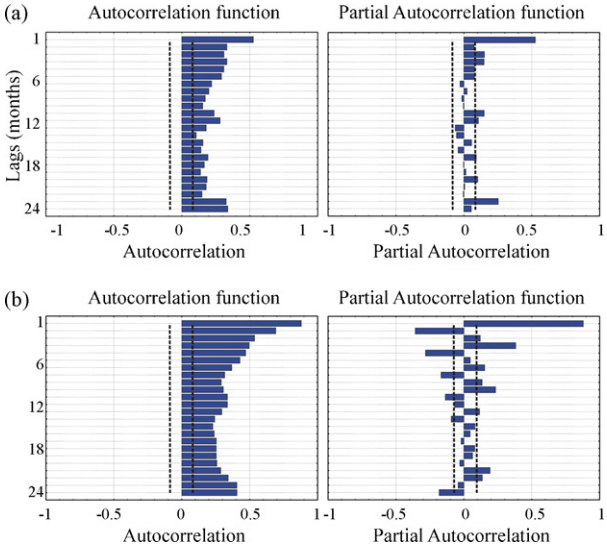

Fig. 1. Autocorrelation and partial autocorrelation functions of: (a) original data series and (b) convoluted data series. Dotted lines indicate statistical signification level ($p = 0.05$).

series. It is possible to observe that the autocorrelation function is a typical function corresponding with a time series with trend. Initially the autocorrelations are significantly higher than zero and gradually decrease to zero in function of the number of lags. In this way, the mean changes with the time which imply that time series is not stationary in time. A similar result is observed for the autocorrelation function of smoothed anchovy catches time series (Fig. 1(b)). On the other hand, a Levene's test showed a lack of variances homogeneity between years in both cases (original series: $F = 6.21$; $p < 0.05$; smoothed series: $F = 10.02$; $p < 0.05$). These results indicate that a previous logarithmic transformation and differentiation process in original and smoothed time series could facilitate the identification of the number of inputs to CNN models.

The ACF and PCAF of the transformed anchovy catches time series (original and smoothed series) are show in Fig. 2. In both cases, a first differentiation ($d = 1$) was enough to eliminate the trend. The PCAF of original series showed significant partial autocorrelation for 1, 2, 3, 4, 5, 11, 16, 19 and 22 lags while the autocorrelation was evident for 1, 2, 10, 12, 14, 22, 23 and 24 lags. When the original data series was smoothed, significant partial autocorrelations were found for 1 to 10 and 12, 14, 15, 17, 18, 20, 21 and 23 lags. In this case, the autocorrelation were significant for 1, 2, 3, 8, 9, 11, 12, 14, 15, 20, 21, 23, and 24 lags. This significant lags configuration indicates that underlies an autoregressive behaviour in the catch series characterized by a short time non-seasonal dependence (approximately 1–6 months) and a medium time seasonal dependence (approximately 18–24 months).

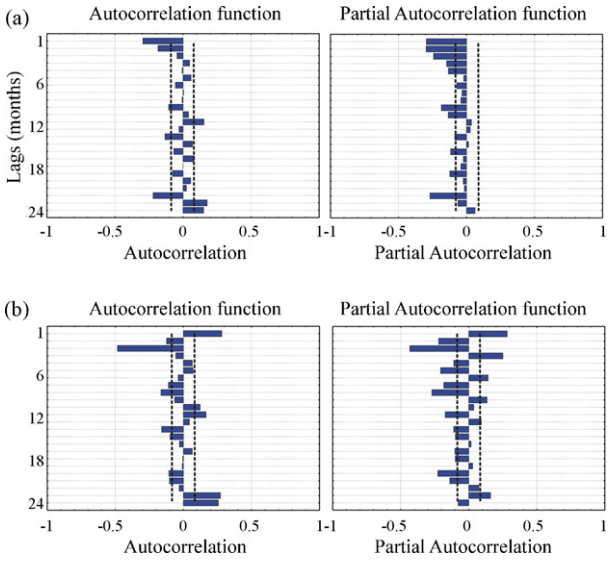

Fig. 2. Autocorrelation and partial autocorrelation functions of differentiated data series ($d = 1$) corresponding with: (a) original data series and (b) convoluted data series. Dotted lines indicate statistical signification level ($p = 0.05$).

The FFT analysis of original data series detected seasonal components of large, medium and short time (Table 1) of spectral density indicating that the large time seasonal component was

Table 1
Spectral analysis of the original series of anchovy catches arranged by spectral density

Frequency (months^{-1})	Period (months)	Period (years)	Cosine coefficient	Sine coefficient	Periodogram	Spectral density
0.0019	516.0	43.0	22432.55	−14907.04	1.87×10^{11}	1.34×10^{11}
0.0039	258.0	21.5	−10911.46	23661.44	1.75×10^{11}	1.31×10^{11}
0.0058	172.0	14.3	5304.40	9609.07	3.11×10^{10}	6.46×10^{10}
0.0329	30.4	2.5	−14810.30	8628.08	7.58×10^{10}	5.12×10^{10}
0.0310	32.3	2.7	4998.30	−13178.40	5.13×101	4.83×10^{10}
0.1725	5.8	0.5	−11973.26	11434.57	7.07×10^{10}	3.70×10^{10}
0.0116	86.0	7.2	−11353.96	3834.66	3.71×10^{10}	3.61×10^{10}
0.0136	73.7	6.1	7959.20	10633.99	4.55×10^{10}	3.19×10^{10}
0.0833	12.0	1.0	4853.53	13101.90	5.04×10^{10}	3.11×10^{10}
0.0349	28.7	2.4	−6700.71	3500.62	1.47×10^{10}	2.99×10^{10}
0.2500	4.0	0.3	−648.27	11908.01	3.67×10^{10}	2.84×10^{10}
0.0097	103.2	8.6	864.30	−11611.45	3.50×10^{10}	2.79×10^{10}
0.0484	20.6	1.7	10618.63	−5713.19	3.75×10^{10}	2.77×10^{10}
0.0291	34.4	2.9	9207.79	4574.93	2.73×10^{10}	2.75×10^{10}
0.2519	4.0	0.3	10303.78	6506.65	3.83×10^{10}	2.69×10^{10}
0.1705	5.9	0.5	−5486.26	−224.92	7.78×10^{9}	2.68×10^{10}
0.0853	11.7	1.0	7992.63	6967.06	2.90×10^{10}	2.58×10^{10}
0.0078	129.0	10.8	3025.62	712.01	2.49×10^{9}	2.46×10^{10}
0.1744	5.7	0.5	4333.00	4561.54	1.02×10^{10}	2.36×10^{10}
0.0504	19.8	1.7	9568.14	−836.68	2.38×10^{10}	2.32×10^{10}
0.0814	12.3	1.0	−3450.83	1972.99	4.08×10^{9}	1.92×10^{10}
0.0465	21.5	1.8	4668.49	7405.36	1.98×10^{10}	1.88×10^{10}
0.0155	64.5	5.4	2825.97	−1526.77	2.66×10^{9}	1.85×10^{10}
0.1686	5.9	0.5	7349.51	6299.98	2.42×10^{10}	1.66×10^{10}
0.2481	4.0	0.3	−5458.41	−3581.71	1.10×10^{10}	1.56×10^{10}
0.0950	10.5	0.9	4780.93	9333.27	2.84×10^{10}	1.56×10^{10}
0.0523	19.1	1.6	3657.52	−5474.28	1.12×10^{10}	1.41×10^{10}
0.0562	17.8	1.5	9759.69	−447.76	2.46×10^{10}	1.37×10^{10}
0.0368	27.2	2.3	−6973.88	1260.86	1.30×10^{10}	1.34×10^{10}

For 29 principal frequencies the period (months and years), cosine coefficient, sine coefficient, periodogram value and spectral density are shown.

controlled by periods of 43, 22 and 14 years, approximately. At medium time scale, 2.5, 3 and 1 years were detected while periods of 6 and 4 months were observed at short time. Similar results were found for smoothed catch series although in this case, the periods corresponding to high frequencies were significantly less important as a consequence of smoothing process.

In this way, non-seasonal and seasonal models were calibrated and externally validated considering the 6 previous months and 1 year as inputs to the neural approximations, respectively. Also, variations around 6 and 12 months were tested with the same calibration and external validation sets.

3.2. Neural networks results: original and smoothed time series

Fig. 3(a) shows the best results in the type I external validation phase of the 120 CNN models calibrated with original catches data series and considering the 6 previous months as inputs (non-seasonal autoregressive CNN). The architecture of the best model, considering for its selection an assessment based on the different accuracy measures, was 6–5*s*–5*s*–1*l*.

The determination coefficient between observed and estimated catches in this validation phase indicated that only a 50.41% of the explained variance was captured by the model. In this case, this level of explained variance implied a standard error prediction (SEP) of 90.26%, a root mean standard error (RMSE) of 42182.6 tonnes and a mean absolute error of 27175.08 tonnes. A detailed analysis of these results showed that the poor performance was mainly due to the occurrence of a displacement between estimated and observed values, as indicated persistence index close to zero (PI = 0.106). These results were only slightly better than the obtained for the naïve model. However, the calculated values for E_2, ARV, AIC and BIC indicated that the forecast of the non-seasonal autoregressive 6–5*s*–5*s*–1*l* CNN model considering original catches was worst than NF1 model prediction (Table 2).

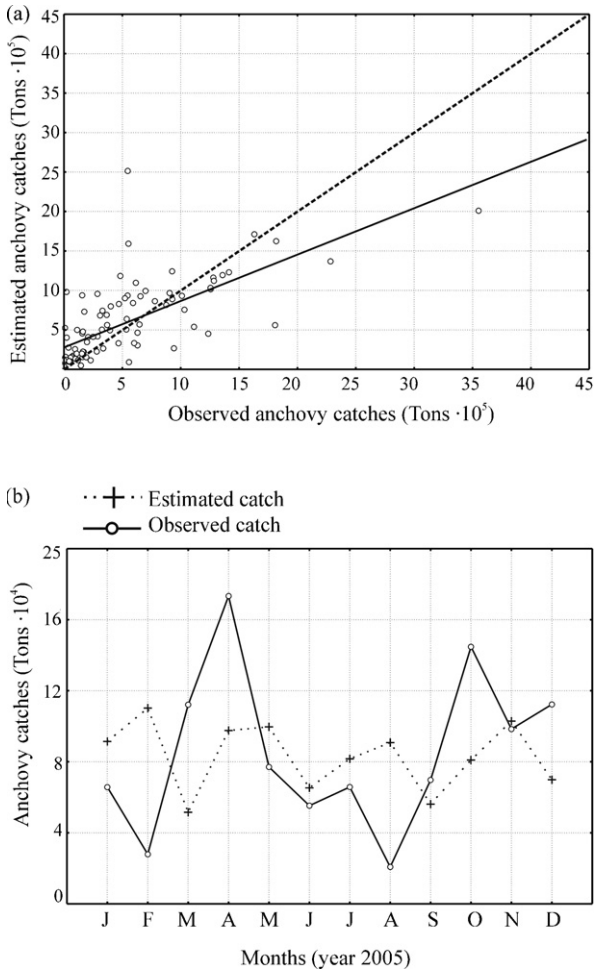

Fig. 3. (a) Scatterplot of observed anchovy catches vs. estimated anchovy catches from 6–5*s*–5*s*–1*l* model calibrated with original data series (EVI; type I external validation; $N = 99$); (b) monthly observed and estimated anchovy catches in the year 2005 (EVII, type II external validation; $N = 12$).

Table 2
Measures of accuracy calculated in the external validation (type I; EVI) for the best non-seasonal and seasonal models calibrated with original and convoluted data (CNN producing the best performance in each model was selected from a pool of 30 repetitions)

Accuracy measures	NF1	6–5*s*–5*s*–1*l*[a]	6–15*s*–15*s*–1*l*[b]	7–5*s*–5*s*–1*l*[c]	7–10*s*–10*s*–1*l*[d]
R	0.5267	0.7100	0.9071	0.4962	0.8875
R^2	0.2773	0.5041	0.8227	0.2462	0.7873
PI	0	0.1056	0.7732	0.2037	0.7629
RMSE (tonnes)	63378.87	42182.27	30400.49	65836.04	30668.54
SEP (%)	111.1082	90.2642	52.1075	100.6922	51.1724
MAE (tonnes)	36825.36	27175.08	18882.79	43494.13	20756.98
E_2	0.0539	−0.9333	0.8117	−0.3188	0.7724
ARV	0.9461	1.9333	0.1883	1.3188	0.2276
m	1	60	330	65	180
AIC	9.6078	10.4624	15.7005	10.8490	12.6097
BIC	9.6091	10.4597	15.6709	10.8464	12.6018

In all cases $N = 99$ selected in the range January 1963–December 2004. NF1 shows the accuracy measures for the basic naïve model.
[a] Inputs in the calibration phase = original data; model type = non-seasonal autoregressive CNN.
[b] Inputs in the calibration phase = convoluted data; model type = non-seasonal autoregressive CNN.
[c] Inputs in the calibration phase = original data; model type = seasonal ($S = 12$ [1 year]; $P = 1$ [1 month]) autoregressive CNN.
[d] Inputs in the calibration phase = convoluted data; model type = seasonal ($S = 12$ [1 year]; $P = 1$ [1 month]) autoregressive CNN.

Table 3
Measures of accuracy calculated in the external validation (type II; EVII) for the best non-seasonal and seasonal models calibrated with original and convoluted data (CNN producing the best performance in each model was selected from a pool of 30 repetitions)

Accuracy measures	NF1	6–5*s*–5*s*–1*l*[a]	6–15*s*–15*s*–1*l*[b]	7–5*s*–5*s*–1*l*[c]	7–10*s*–10*s*–1*l*[d]
R	0.5267	−0.1236	0.7481	−0.0146	0.7543
R^2	0.2773	0.0152	0.5597	0.0002	0.5690
PI	0	0.1807	0.6875	0.2423	0.6868
RMSE (tonnes)	63378.87	48883.71	30191.77	47008.45	30222.78
SEP (%)	111.1082	57.3290	35.4078	55.1297	35.4441
MAE (tonnes)	36825.36	40992.84	23477.92	37992.29	23666.10
E_2	0.0539	0.1158	0.6572	−0.1870	0.5093
ARV	0.9461	0.8842	0.3428	1.1870	0.4907
m	1	60	330	65	180
AIC	9.6078	19.3783	63.9598	20.1776	38.9607
BIC	9.6091	14.7742	38.6373	15.1899	25.1483

In all cases $N = 12$ corresponding with year 2005. NF1 shows the accuracy measures for the basic naïve model.
[a] Inputs in the calibration phase = original data; model type = non-seasonal autoregressive CNN.
[b] Inputs in the calibration phase = convoluted data; model type = non-seasonal autoregressive CNN.
[c] Inputs in the calibration phase = original data; model type = seasonal ($S = 12$ [1 year]; $P = 1$ [1 month]) autoregressive CNN.
[d] Inputs in the calibration phase = convoluted data; model type = seasonal ($S = 12$ [1 year]; $P = 1$ [1 month]) autoregressive CNN.

The insufficient forecast capacity of this model was also observed in type II external validation phase (year 2005). In this case, the model tends to estimate anchovy catches around the mean value of observed anchovy catches for 2005 (Fig. 3(b)). This behaviour implied a very low level of explained variance ($R^2 = 0.0152$), a very high mean absolute error (MAE = 40992.84) and AIC and BIC values higher than those calculated for the NF1 model. In spite of it, the rest of accuracy measures provided values slightly better than the obtained with the naïve model (Table 3).

Significantly better results were obtained when the inputs to the models were convoluted by means a smoothed function. In this case, the architecture which provided the best prediction had 15 neurons in the first and second hidden layers (6–15*s*–15*s*–1*l*). For this model, all accuracy measures except AIC and BIC were better than those calculated for NF1 in both external validation phases (Tables 2 and 3). Compared with the best neural network calibrated with non-smoothed data (6–5*s*–5*s*–1*l* model), the 6–15*s*–15*s*–1*l* model improved significantly all error terms in the type II external validation phase (Table 3). In type I external validation phase, only AIC and BIC were better for the 6–5*s*–5*s*–1*l* model (Table 2).

Fig. 4(a) shows the regression between observed and estimated anchovy catches in type I external validation phase for the best neural network calibrated with smoothed data (6–15*s*–15*s*–1*l* model). It is possible to observe a good fit of data to line 1:1, although exists a higher dispersion level above 200,000 tonnes. This is a consequence that, even thought the model captured the general behaviour of data series, tended to undervalue some pick catches as March, April and October of the year 2005 (type II external validation phase) (Fig. 4(b)).

Very similar results were found when the seasonal autoregressive CNN models were used. In this case, the best results were obtained for a seasonal order of 12 months ($S = 12$) and a seasonal autoregressive term $P = 1$. The effect of the seasonality was common to CNNs calibrated with original and convoluted data, although the hidden neural architecture which

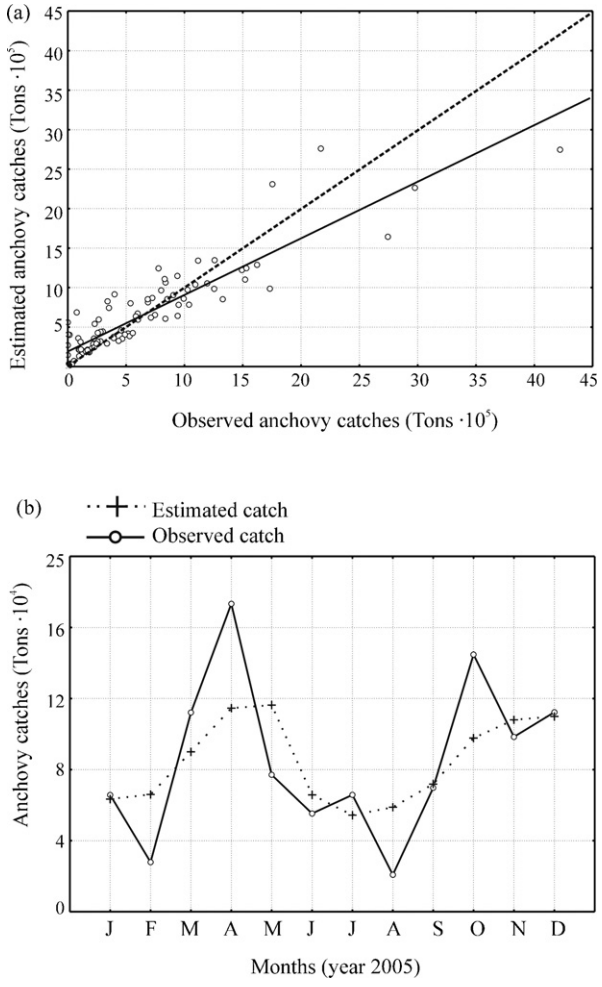

Fig. 4. (a) Scatterplot of observed anchovy catches vs. estimated anchovy catches from 6–15*s*–15*s*–1*l* model calibrated with convoluted data series (EVI; type I external validation; $N = 99$); (b) monthly observed and estimated anchovy catches in the year 2005 (EVII, type II external validation; $N = 12$).

Table 4
Accuracy measures of the best recurrent CNN models for each hidden neuronal architecture in the external validation (type I; EVI) (CNN producing the best performance in each model was selected from a pool of 30 repetitions)

Hidden neuronal architecture	Accuracy measures									
	R	R^2	RMSE (tonnes)	SEP (%)	E_2	ARV	MAE (tonnes)	PI	AIC	BIC
NFI	0.5267	0.2773	63378.9	111.1	36825	0.0539	0.9461	0	9.6078	9.6078
5	0.7012	0.4917	72805.44	122.9407	−0.4422	1.4422	69870.64	0.0734	8.4358[a]	8.4342[a]
10	0.6865	0.4713	48117.96	81.2523	0.3700	0.6300	46530.72	0.2565	8.7831	8.7800
15	0.8287	0.6867	35237.91	59.5034	0.6621	0.3379	34280.63	0.6284	9.2196	9.2150
20	0.8349	0.6971	35380.36	59.7439	0.6594	0.3406	26540.22	0.6692	9.9302	9.9240
25	0.7300	0.5329	41866.48	70.6965	0.5230	0.4769	55698.81	0.3377	10.7835	10.7757
30	0.7704	0.5935	31327.18[a]	46.1452[a]	0.7968[a]	0.2032[a]	23200.82[a]	0.6547	11.1200	11.1107
35	0.9222	0.8505	35966.96	60.7345	0.6480	0.3520	28432.45	0.7324[a]	12.0657	12.0549
40	0.9383[a]	0.8804[a]	32983.68	55.6969	0.7039	0.2960	27030.68	0.7241	12.6975	12.6852
45	0.6471	0.4187	86062.63	145.3270	−1.0152	2.0153	71690.46	0.1129	14.2376	14.2237
50	0.5600	0.3136	90763.78	153.2655	−1.2414	2.2415	81640.52	0.1238	14.9909	14.9755
55	0.8433	0.7112	111621.66	188.4864	−2.3900	3.3900	86060.41	0.1697	15.8776	15.8607

NF1 shows the accuracy measures for the basic naïve model.
[a] Best results.

provided the best results was different (7–5*s*–5*s*–1*l* for original data and 7–10*s*–10*s*–1*l* for convoluted data). The external validation phases (type I and II) for the 7–5*s*–5*s*–1*l* model showed accuracy measures very close to 6–5*s*–5*s*–1*l* model calibrated with original data. Type I external validation of the best CNN calibrated with convoluted data provided worse values (except AIC and BIC) than those calculated for the non-seasonal model (Table 2). On the other hand, the values obtained in the type II external validation phase were slightly better (Table 3).

3.3. Recurrent neural networks results

Table 4 shows the best prediction results reported for the recurrent CNNs in the type I external validation phase. In this case, the architectures which provided the best results (based on a global evaluation of all accuracy measures) had among 30 and 40 neurons in the hidden layer. This way, the 6–40*s*–1 model explained 88% of the variance, but the lowest error levels (RMSE = 31327.18 tonnes; SEP = 46.1452%) and displacement (PI = 0.7324) between observed and estimated data were found for 30 and 35 neurons in the hidden layer, respectively.

Globally, it was possible to observe that the recurrent CNNs predicted better than the classic autoregressive CNN models calibrated with convoluted and non-convoluted data, but did not overcome the forecast capacity of CNN + ARIMA autoregressive hybrid model calibrated with convoluted data (Section 3.4). Also, this trend was observed in the type II external validation phase.

Table 5
Measures of accuracy calculated in the external validation (type I; EVI) for the best non-seasonal and seasonal models calibrated with convoluted data and hybridised with ARIMA models

Accuracy measures	NF1	6–15*s*–15*s*–1*l*[a] + ARIMA(2,0,0)	7–10*s*–10*s*–1*l*[b] + ARIMA(1,0,1)(1,0,1)$^{S=12}$
R	0.5267	0.9175	0.9030
R^2	0.2773	0.8418	0.8155
PI	0	0.7778	0.7874
RMSE (tonnes)	63378.87	30090.76	29478.51
SEP (%)	111.1082	51.5766	49.7780
MAE (tonnes)	36825.36	17482.61	21616.97
E_2	0.0539	0.8153	0.7915
ARV	0.9461	0.1847	0.2084
m	1	332[c]	184[d]
AIC	9.6078	15.7324	12.7723
BIC	9.6091	15.7027	12.7384
ARIMA parameters[e]		$\phi_1 = -0.2514$; $\phi_2 = -0.2180$	$\phi_1 = 0.3764$; $\theta_1 = 0.6150$; $\Phi_1 = 0.8779$; $\Theta_1 = 0.7344$

In all cases $N = 99$ selected in the range January 1963–December 2004. NF1 shows the accuracy measures for the basic naïve model.
[a] Inputs in the calibration phase = convoluted data; model type = non-seasonal autoregressive CNN.
[b] Inputs in the calibration phase = original data; model type = seasonal ($S = 12$ [1 year]; $P = 1$ [1 month]) autoregressive CNN.
[c] $m = 330$ CNN weights + 2 autoregressive parameters of ARIMA model.
[d] $m = 184$ CNN weights + 1 non-seasonal autoregressive parameter + 1 non-seasonal moving average parameter + 1 seasonal autoregressive parameter + 1 seasonal moving average parameter of ARIMA model.
[e] All parameters $p < 0.05$.

Table 6
Measures of accuracy calculated in the external validation (type II; EVII) for the best non-seasonal and seasonal models calibrated with convoluted data and hybridised with ARIMA models

Accuracy measures	NF1	6–15*s*–15*s*–1*l*[a] + ARIMA(2,0,0)	7–10*s*–10*s*–1*l*[b] + ARIMA(1,0,1)(1,0,1)$^{S=12}$
R	0.5267	0.9344	0.8380
R^2	0.2773	0.8731	0.7022
PI	0	0.7703	0.7459
RMSE (tonnes)	63378.87	25881.32	27221.58
SEP (%)	111.1082	30.3527	31.9245
MAE (tonnes)	36825.36	19571.75	21122.99
E_2	0.0539	0.6001	0.6019
ARV	0.9461	0.3598	0.3981
m	1	332[c]	184[d]
AIC	9.6078	64.1593	39.5365
BIC	9.6091	38.6833	25.4173
ARIMA parameters[e]		$\phi_1 = -0.2514$; $\phi_2 = -0.2180$	$\phi_1 = 0.3764$; $\theta_1 = 0.6150$; $\Phi_1 = 0.8779$; $\Theta_1 = 0.7344$

In all cases $N = 12$ corresponding with year 2005. NF1 shows the accuracy measures for the basic naïve model.
[a] Inputs in the calibration phase = convoluted data; model type = non-seasonal autoregressive CNN.
[b] Inputs in the calibration phase = original data; model type = seasonal ($S = 12$ [1 year]; $P = 1$ [1 month]) autoregressive CNN.
[c] $m = 330$ CNN weights + 2 autoregressive parameters of ARIMA model.
[d] $m = 184$ CNN weights + 1 non-seasonal autoregressive parameter + 1 non-seasonal moving average parameter + 1 seasonal autoregressive parameter + 1 seasonal moving average parameter of ARIMA model.
[e] All parameters $p < 0.05$.

3.4. Hybrid model results

Significant improvements were found when the best CNN models were calibrated together with ARIMA models. In case of the non-seasonal model (6–15*s*–15*s*–1*l*) initially calibrated with convoluted data, the best estimate was obtained by applying an ARIMA(2,0,0) to residuals of the CNN model (Table 5). For this configuration better accuracy measures and parameters ϕ_i with a level of acceptable statistical significance ($p < 0.05$) were achieved. This way, in the type I external validation phase, the hybrid model improved all accuracy measures (except AIC and BIC).

Also, the forecast capacity of this model was shown in the type II external validation phase where the explained variance reached a level of 87.31% and the standard error of prediction was slightly above 30% (the lowest level among all calibrated and validated models) (Table 6). This is related with a low dispersion between observed and estimated data along the line 1:1 and a low displacement between both time series (Fig. 5).

Likewise good estimations were reached when the best seasonal CNN model was hybridised with an ARIMA model. In this case, the ARIMA configuration which provided the best results had seasonal and non-seasonal parameters with a seasonal cycle of 12 months ($S = 12$). In type I external validation phase, this model explained a higher level of variance at the expense of reducing the lag between observed and estimated values and using a more complex configuration. This way, the persistence index was maximum for this model (PI = 0.7874), but the absolute average error was higher than the calculated for 6–15*s*–15*s*–1*l* + ARIMA(2,0,0). Also, this effect was observed in type II external validation phase (Tables 5 and 6).

Fig. 6(a) shows the schematic representation of the forecasted anchovy catches as a function of the observed catches for the type I external validation phase from the best seasonal hybrid model. It is possible to observe a good fit of the scatter data along the line 1:1, although as same as the rest of the CNNs calibrated, this model tended to undervalue high catch levels and in a less important form overvalue low catch levels. In spite of this, the model captured the general trend of the anchovy catches data series (Fig. 6(b)).

4. Discussion

The adequacy of computational neural networks models for monthly anchovy catches forecasting in north area of Chile was analysed in this paper. The best correlation and error statistics were obtained based on hybrid configuration CNN + ARIMA having as input data the anchovy catches in 6 previous months combined with a convolution process of the input sequence by mean a smooth function. The results thus achieved were better than those obtained with typical autoregressive training based only on catches data at various time delays. Also, the results were better than those obtained with recurrent neural networks (Elman model) and convoluted autoregressive CNN models.

As suggested Legates and McCabe (1999), a multicriteria performance assessment based on different accuracy measures was appropriated to select the best models. In some cases, the explained variances were significantly high pointed towards the good performance of the model, but the values of RMSE, SEP, E_2, ARV, MAE and/or PI were significantly worse than those obtained with others models and even with the naïve model (NFI). This way, high correlations can be achieved by mediocre or poor models. Similar conclusions were obtained in forecasting of different kinds of time variables (Garrick et al., 1978; Willmott et al., 1985; Stergiou et al., 1997; Pulido-Calvo and Portela, 2007; Velo-Suárez and Gutiérrez-Estrada, 2007).

In typical autoregressive CNN, convoluted autoregressive CNN and recurrent CNN models, the principal source of error

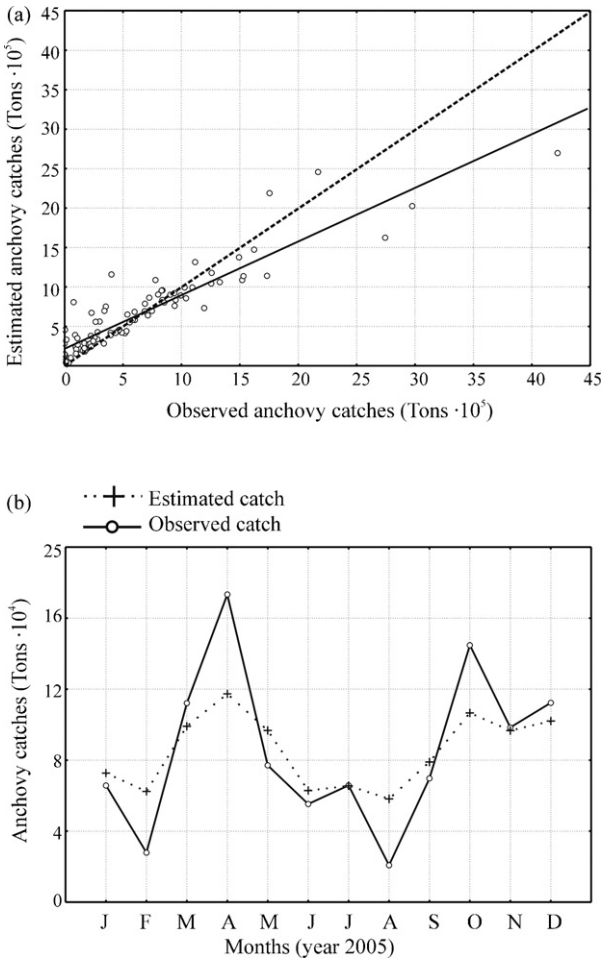

Fig. 5. (a) Scatterplot of observed anchovy catches vs. estimated anchovy catches from 6–15*s*–15*s*–1*l* + ARIMA(2,0,0) model calibrated with convoluted data series (EVI; type I external validation; $N = 99$); (b) monthly observed and estimated anchovy catches in the year 2005 (EVII, type II external validation; $N = 12$).

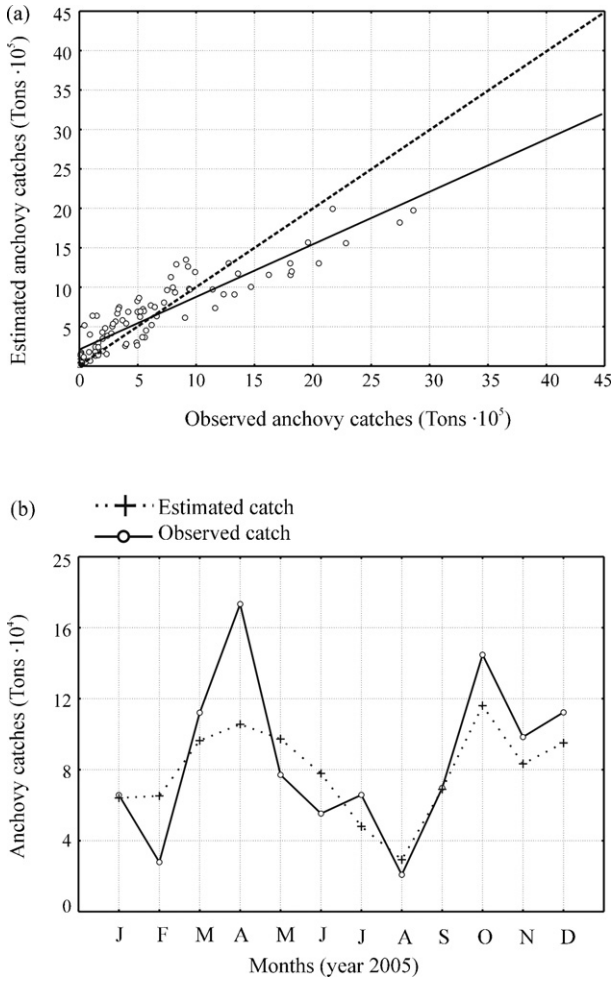

Fig. 6. (a) Scatterplot of observed anchovy catches vs. estimated anchovy catches from 7–10*s*–10*s*–1*l* + ARIMA(1,0,1)(1,0,1)$^{S=12}$ model calibrated with convoluted data series (EVI; type I external validation; $N = 99$); (b) monthly observed and estimated anchovy catches in the year 2005 (EVII, type II external validation; $N = 12$).

was found in the displacement of the estimated curves with regard to the observed ones. This effect occurred despite memory was added to network by mean of a buffer containing recent inputs (catches in previous months) which was probably due to the fact that the correlation between observed catches in any consecutive months were most of the time high ($R = 0.53$; $p < 0.05$) and so, each month occurrence is highly responsible for the next monthly realization. On the other hand it is known that a trained feed forward neural network has a static architecture and the output is solely determined by the present input state to the network and not by the initial and past states of the neurons in the network (Pulido-Calvo and Portela, 2007). Park (1998), Abrahart and See (2000), Pulido-Calvo et al. (2003), Gutiérrez-Estrada et al. (2004) and Velo-Suárez and Gutiérrez-Estrada (2007) also reported this 'lag-one' difference between observed and estimated values resulting from multiple regression, ARIMA and CNNs models applied to forecasting of different time variables.

In general, in spite of the wide spectrum of different configurations used in this work, the best results for each configuration type indicated that the models captured the anchovy fishery behaviour. Similar results are reported for other small pelagic fisheries using linear univariate models with long and stable data series. Stergiou et al. (1997) indicated that ARIMA(1,0,1)(0,1,2)^{12} model fitted and forecasted the monthly catches of anchovy (*Engraulis encrasicolus*) in Hellenic waters. In this case, the model provided very high values of R^2 and low values of BIC, medium standard error and mean absolute percentage error. These results contrasted with those obtained from medium-short and non-stable data series by Lloret et al. (2000) which reported that ARIMA models had a low forecast capacity for small pelagic fishes (as *E. encrasicolus* or small *Sardina pilchardus*) in north-western Mediterranean sea. This indicates that in these conditions ARIMA models cannot extract the linear component of data series. This way, CNN models which have a great capacity to mapping highly non-linear relationships between variables may be more appropriate when unstable data series (as anchovy catches in north area of Chile) are used.

In spite of high generalisation capacity that by themselves have the CNN models, the best results in both external validation phases were obtained by CNN + ARIMA hybrid models. In this case, the CNN alone was not able map the linear and

non-linear relationships between variables, which indicates the high complexity of the modeled system. Oceanographic and climatic shifts have been related with marine ecosystem changes for several spatial-temporal scales (Hare et al., 2000; Alheit and Ñiquen, 2004). Specifically, Cañón (1986), Yáñez et al. (1986) and Montecinos et al. (2003) report that the ecosystem of anchovy fishery is conditioned by El Niño events and long-term environmental changes. Generally, not very intensive El Niño events cause horizontal and vertical movements of this fishery resource. Nevertheless, these events are related with environmental shifts of large scale, which is the reason why it is very difficult to identify their effects on anchovy resource as well as to evaluate the interaction between biological processes and the exploitation of this resource. Also, the ecosystem complexity is present at seasonal scale, because water temperature variation and upwelling phenomena in the coast regulate the ecosystem productivity, migration and growth capacity of anchovy stock (Myers et al., 1995; Hutching, 2000; Yáñez et al., 2001).

The results show that independently of the model type and number of hidden neurons, the number of inputs which explained the highest levels of variance in the external validation phase was 6 months. Thus, the 6 previous months were enough to explain a variance level between 84% and 87% when the CNN was hybridised with an ARIMA(2,0,0) model. This can be related with biological aspects of anchovy which have a great influence on fishery features and consequently on the topology of the models. Cubillos et al. (2002) report that among the most important fishery and biological aspects of anchovy are: (a) short life span; (b) fast growth in length, with seasonally oscillating growth rate; (c) spawning time in the winter time (July–August–September); (d) seasonal fishery, with higher catches and fishing effort during January–March every year. These features can be implicit in the configuration of the model. This way, the anchovy recruitment (which occurs approximately in January), is highly non-linear dependent of the spawning because the eggs and larval survival are related with the turbulence index variation and Ekman transport, which ones are related with the direction and wind intensity (Bakum et al., 1974; Santander and Flores, 1983; Parrish et al., 1983; Yáñez et al., 2001). On the other hand, the variance explained by the linear component of the model [ARIMA(2,0,0)], can be associated to spawning secondary process (Castillo et al., 2002) that occurs in November, December and January (approximately 2 months before the highest annual catches).

The good results obtained could indicate that, from the point of view of univariate approach, a *quasi*-complete characterization of anchovy catches has been reached. However, it is also necessary to point out the model limitations (its forecasting capacity and/or it biological significance) in the context of a short-medium term time period. Any method based on univariate techniques assumes that exists some dependence (linear o non-linear) between the time series data. In this way, each observation can be explained as a linear/non-linear function of its past values. This implies that the variance of data series is equivalent to the variances sum of any external variable that has influence on anchovy catches. Therefore, the univariate model that explains shifts of historical data has the capacity to detect the external variables influence implicit in the variance of time series data. However, if new external variables intervene in the system or significant changes in the relationships established by the model are carried out then the forecasting accuracy of the model will be damaged. These changes will imply that new anchovy catch patterns must be considered and therefore the model should be calibrated and validated again.

Acknowledgements

The authors wish to express their gratitude to the AECI (Agencia Española de Cooperación Internacional) for financing this research under Project A/3618/05. We are also grateful to Alejandra Órdenes, Inés Guerrero and Francisco Plaza for their assistance in the anchovy catches data collection. Also, we wish to express our gratitude to Lorenzo Lidueña González for assistance with the critical revision of the English language.

References

Abrahart, R.J., See, L., 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrol. Process. 14, 2157–2172.

Alheit, J., Ñiquen, M., 2004. Regime shift in the Humboldt Current ecosystem. Prog. Oceanogr. 60, 201–222.

Anctil, F., Rat, A., 2005. Evaluation of neural network streamflow forecasting on 47 watersheds. J. Hydrol. Eng. 10 (1), 85–88.

Bakum, A., McLain, D.R., Mayo, F.V., 1974. The mean annual cycle of coastal upwelling off western North America as observed from surface measurements. Fish. Bull. 72 (3), 843–844.

Belgrano, A., Malmgren, B.A., Lindahl, O., 2001. Application of artificial neural networks (ANN) to primary production time-series data. J. Plankton Res. 23 (6), 651–658.

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control. Holden-Day, Oakland, California.

Cañón, J.R., 1986. Distribución de anchoveta (*Engraulis ringens*) en el norte de Chile en relación a determinadas condiciones oceanográficas. Investigación Pesquera 30, 1–122.

Castillo, J., Córdova, J., Saavedra, A., Espejo, M., Gálvez, P., Barbieri, M.A., 2002. Evaluación acústica de la biomasa, abundancia, distribución especial y caracterización de cardúmenes de anchoveta en el período de reclutamiento. Primavera 2001. Evaluación del reclutamiento de anchoveta en la I y II Regiones, temporada 2001–2002. Informes Técnicos FIP-IT/2001-11, 207 pp.

Chen, D.G., Ware, D.M., 1999. A neural network model form forecasting fish stock recruitment. Can. J. Fish. Aquat. Sci. 56, 2385–2396.

Chen, D.G., Hargreaves, N.B., Ware, D.M., Liu, Y., 2000. A fuzzy logic model with genetic algorithm for analysing fish stock-recruitment relationships. Can. J. Fish. Aquat. Sci. 57 (9), 1878–1887.

Chen, D.G., Hare, S.R., 2006. Neural network and fuzzy logic models for pacific halibut recruitment analysis. Ecol. Model. 195, 11–19.

Cubillos, L.A., Bucarey, D.A., Canales, M., 2002. Monthly abundance estimation for common sardine *Strangomera bentincki* and anchovy *Engraulis ringens* in the central-southern area off Chile (34–40°S). Fish. Res. 57, 117–130.

De Vries, B., Principe, J.C., 1991. A theory for neural networks with time delays. Advances in Neural Information Processing Systems, vol. 3. Morgan Kaufmann Publishers, California.

Elman, J.L., 1990. Finding structure in time. Cognitive Sci. 14, 179–211.

Fréon, P., Mullon, C., Voisin, B., 2003. Investigating remote synchronous patterns in fisheries. Fish. Oceanogr. 12 (4/5), 443–457.

Garrick, M., Cunnane, C., Nash, J.E., 1978. A criterion of efficiency for rainfall-runoff models. J. Hydrol. 36 (3/4), 375–381.

Griñó, R., 1992. Neural networks for univariate time series forecasting and their application to water demand prediction. Neural Network World 2 (5), 437–450.

Gutiérrez-Estrada, J.C., de Pedro-Sanz, E., López-Luque, R., Pulido-Calvo, I., 2004. Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (*Anguilla anguilla* L.) intensive rearing system. Aquacult. Eng. 31, 183–203.

Hansen, J.V., Nelson, R.D., 1997. Neural networks and traditional time series methods: a synergistic combination in state economic forecasts. IEEE Trans. Neural Networks 8 (4), 863–873.

Hardman-Mountford, N.J., Richardson, A.J., Boyer, D.C., Kreiner, A., Boyer, H.J., 2003. Relating sardine recruitment in the North Benguela to satellite-derived sea surface height using a neural network pattern recognition approach. Prog. Oceanogr. 59, 241–255.

Hare, S.R., Minobe, S., Wooster, W.S., McKinnell, S., 2000. An introduction to the PICES symposium on the nature and impacts of North Pacific climate regime shifts. Prog. Oceanogr. 47, 99–102.

Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York.

Holton-Wilson, J., Keating, B., 1996. Previsiones en los negocios. IRWIN, Madrid.

Huse, G., Gjosaeter, H., 1999. A neural network approach for predicting stock abundance of the Barents Sea capelin. Sarsia 84 (5/6), 457–464.

Huse, G., Ottersen, G., 2003. Forecasting recruitment and stock biomass of Northeast Artic cod using neural networks. Sci. Mar. 67, 325–335.

Hutching, J.A., 2000. Collapse and recovery of marine fishes. Nature 406, 882–885.

Hyun, K., Song, M.Y., Kim, S., Chon, T.S., 2005. Using an artificial neural network to patternize long-term fisheries data from South Korea. Aquat. Sci. 67, 382–389.

Iyer, M.S., Rhinehart, R.R., 1999. A method to determine the required number of neural-network training repetitions. IEEE Trans. Neural Networks 10 (2), 427–432.

Kitanidis, P.K., Bras, R.L., 1980. Real time forecasting with a conceptual hydrological model. 2. Applications and results. Water Resour. Res. 16 (6), 1034–1044.

Komatsu, T., Aoki, I., Mitani, I., Ishii, T., 1994. Prediction of the catch of japanese sardine larvae in sagami bay using a neural-network. Fish. Sci. 60 (4), 385–391.

Lassen, H., Medley, P., 2001. Virtual Population Analysis. A practical manual for stock assessment. FAO Fish. Tech. Paper 400.

Legates, D.R., McCabe Jr., G.J., 1999. Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35 (1), 233–241.

Lloret, J., Lleonart, J., Solé, I., 2000. Time series modelling of landing in Northwest Mediterranean Sea. ICES J. Mar. Sci. 57, 171–184.

Makridakis, S., Wheelwrighy, S., McGee, V., 1983. Forecasting: Methods and Applications. John Wiley and Sons, New York.

Maravelias, C.D., Haralaobus, J., Papaconstantinou, C., 2003. Predicting demersal fish species distributions in the Mediterranean Sea using artificial neural networks. Mar. Ecol. Prog. Ser. 255, 249–258.

McLaughlin, R.L., 1983. Forecasting models: sophisticated or naïve? J. Forecasting 2 (3), 274–276.

Montecinos, A., Purca, S., Pizarro, O., 2003. Interannual-to-interdecadal sea surface temperature variability along the westerns coast of South America. Geophys. Res. Lett. 30, 1570.

Myers, R.A., Barrowman, N.J., Hutching, J.A., Rosemberg, A.A., 1995. Population dynamics of exploited fish stocks at low populations levels. Science 269, 1106–1108.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. I. A discussion of principles. J. Hydrol. 10, 282–290.

Park, H.H., 1998. Analysis and prediction of walleye pollock (*Theragrana chalcogramma*) landings in Korea by time series analysis. Fish. Res. 38, 1–7.

Parrish, R.H., Bakum, A., Husby, D.M., Nelson, C.S., 1983. Comparative climatology of selected environmental processes in relation to Eastern boundary current pelagic fish reproduction. FAO Fish. Rep. 291 (3), 731–777.

Prybutok, V.R., Yi, J., Mitchell, D., 2000. Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations. Eur. J. Oper. Res. 122, 31–40.

Pulido-Calvo, I., Roldán, J., López-Luque, R., Gutiérrez-Estrada, J.C., 2003. Demand forecasting for irrigation water distribution system. J. Irrig. Drain. Eng. 129 (6), 422–431.

Pulido-Calvo, I., Portela, M.M., 2007. Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. J. Hydrol. 332 (1/2), 1–15.

Qi, M., Zhang, G.P., 2001. An investigation of model selection criteria for neural network time series forecasting. Eur. J. Oper. Res. 132, 666–680.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. 'Learning' representations by backpropagation errors. Nature 323, 533–536.

SAG, 1963–1977. Anuarios Estadísticos de Pesca. Servicio Agrícola y Ganadero. Ministerio de Agricultura, Chile.

Santander, H., Flores, R., 1983. Los desoves y distribución larval de cuatro especies pelágicas y sus relaciones con las variaciones del medio ambiente marino frente al Perú. FAO Inf. Pesca 292 (3), 835–867.

SERNAPESCA, 1978-2004. Anuarios Estadísticos de Pesca. Servicio Nacional de Pesca. Ministerio de Economía, Fomento y Reconstrucción, Chile.

Shepherd, A.J., 1997. Second-order Methods for Neural Networks. Springer, New York.

Shepherd, J.G., 1999. Extended survivors analysis: an improved method for analysis of catch-at-age data and abundance indices. ICES J. Mar. Sci. 56, 584–591.

Stergiou, K.I., Christou, E.D., Petrakis, G., 1997. Modelling and forecasting monthly fisheries catches: comparison of regression, univariate and multivariate time series methods. Fish. Res. 29, 55–95.

Tan, Y., van Cauwenberghe, A., 1999. Neural-network-based *d*-step-ahead predictors for nonlinear systems with time delay. Eng. Appl. Artif. Intell. 12 (1), 21–25.

Tsoukalas, L.H., Uhrig, R.E., 1997. Fuzzy and Neural Approaches in Engineering. Wiley Interscience, New York.

Velo-Suárez, L., Gutiérrez-Estrada, J.C., 2007. Artificial neural network approaches to one-step weekly prediction of *Dinophysis acuminata* blooms in Huelva (Western Andalucía, Spain). Harmful Algae 6, 361–371.

Ventura, S., Silva, M., Pérez-Bendito, D., Hervás, C., 1995. Artificial neural networks for estimation of kinetic analytical parameters. Anal. Chem. 67 (9), 1521–1525.

Wedding II, D.K., Cios, K.J., 1996. Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. Neurocomputing 10, 149–168.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res. 90, 8995–9005.

Yáñez, E., Barbieri, M.A., Barra, O., 1986. Evaluación de los principales recursos pelágicos explotados en la zona norte de Chile entre 1957 y 1985. In: Arana, P. (Ed.), La pesca en Chile. Escuela de Ciencias del Mar. UCV, Valparaíso, pp. 183–194.

Yáñez, E., Barbieri, M.A., Silva, C., Nieto, K., Espíndola, F., 2001. Climate variability and pelagic fisheries in northern Chile. Prog. Oceanogr. 49, 581–596.

Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. J. Hydrol. 181, 23–48.

Zeng, Z., Yan, H., Fu, A.M.N., 2001. Time-series prediction based on pattern classification. Artif. Intell. Eng. 15, 61–69.

Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50, 159–175.